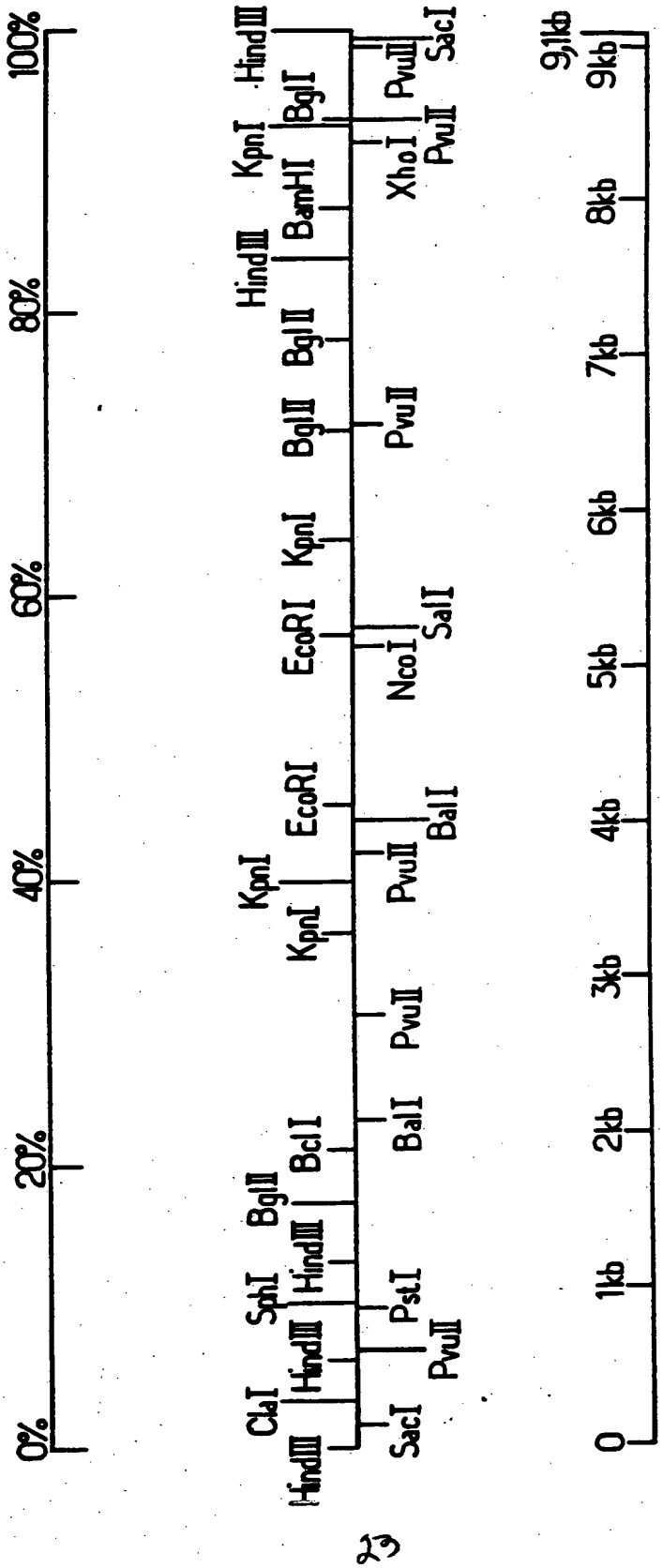
FIG.1.


0%
20%
40%
60%
80%
100%
0
1kb
2kb
3kb
4kb
5kb
6kb
7kb
8kb
9kb
9,1kb
HindIII
ClaI
HindIII
SacI
PvuII
PstI
SphI
HindIII
BglII
BclI
BalI
PvuII
KpnI
KpnI
EcoRI
PvuII
BalI
NcoI
SalI
EcoRI
KpnI
BglII
PvuII
HindIII
BglII
BamHI
KpnI
BglI
PvuII
XhoI
PvuII
HindIII
SacI

23

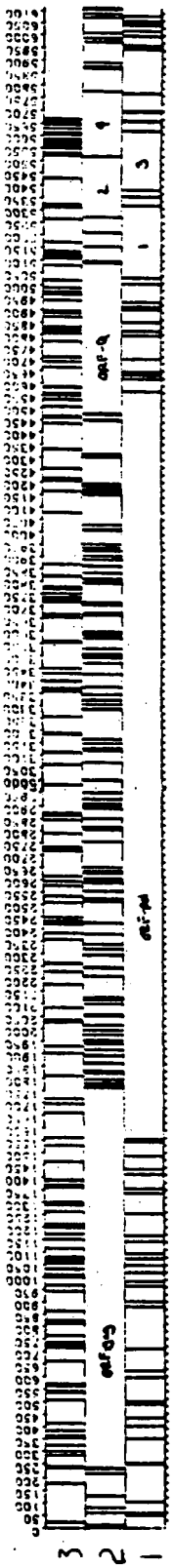


Fig. 2

771248

08/177920

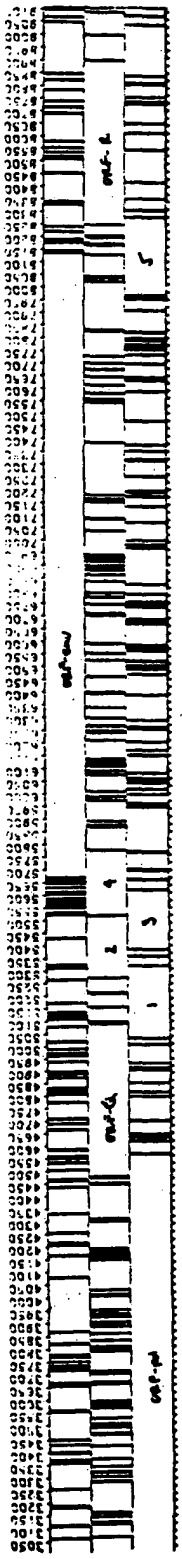


Fig. 3

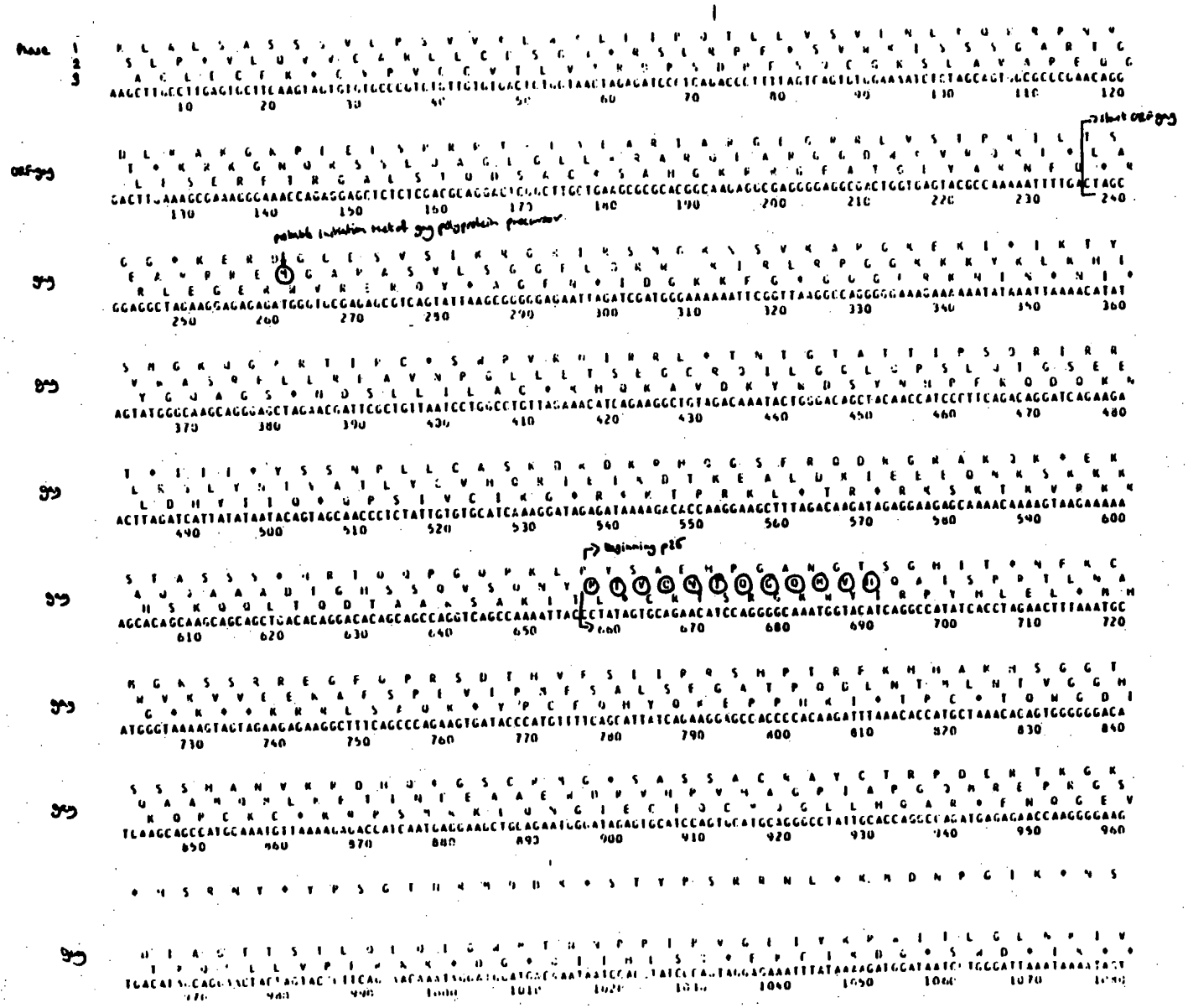


Fig 4

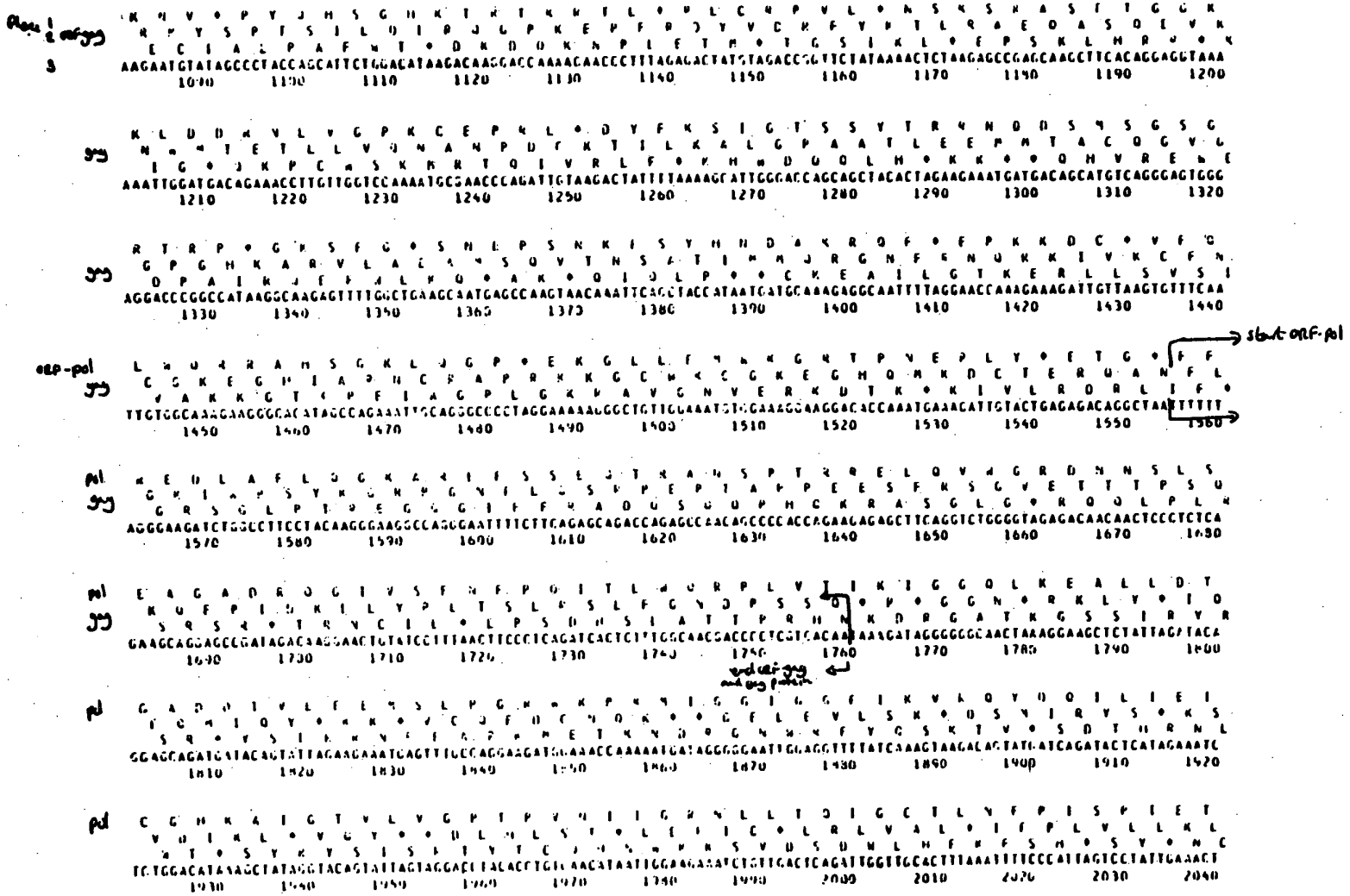


Fig. 5

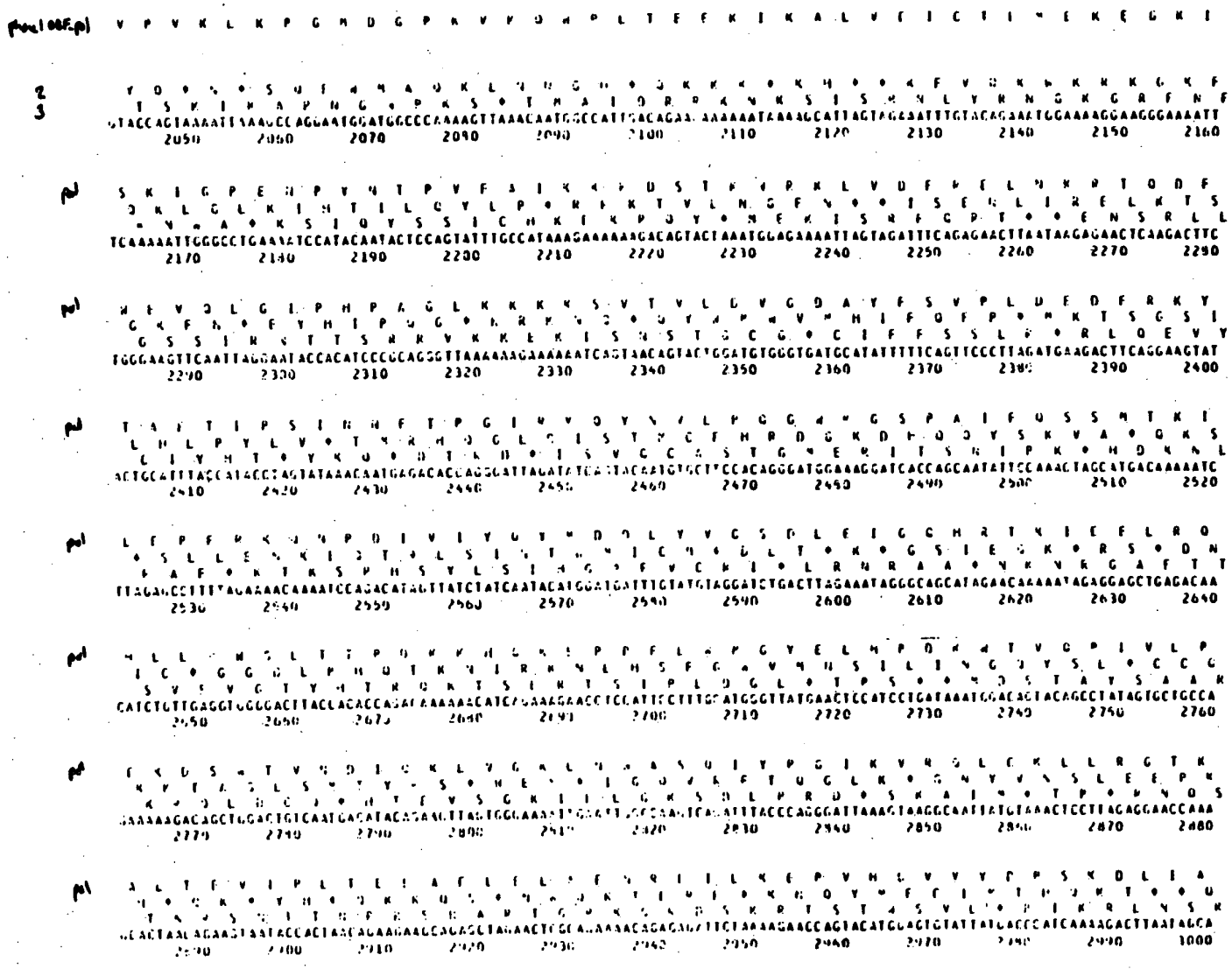


Fig 6

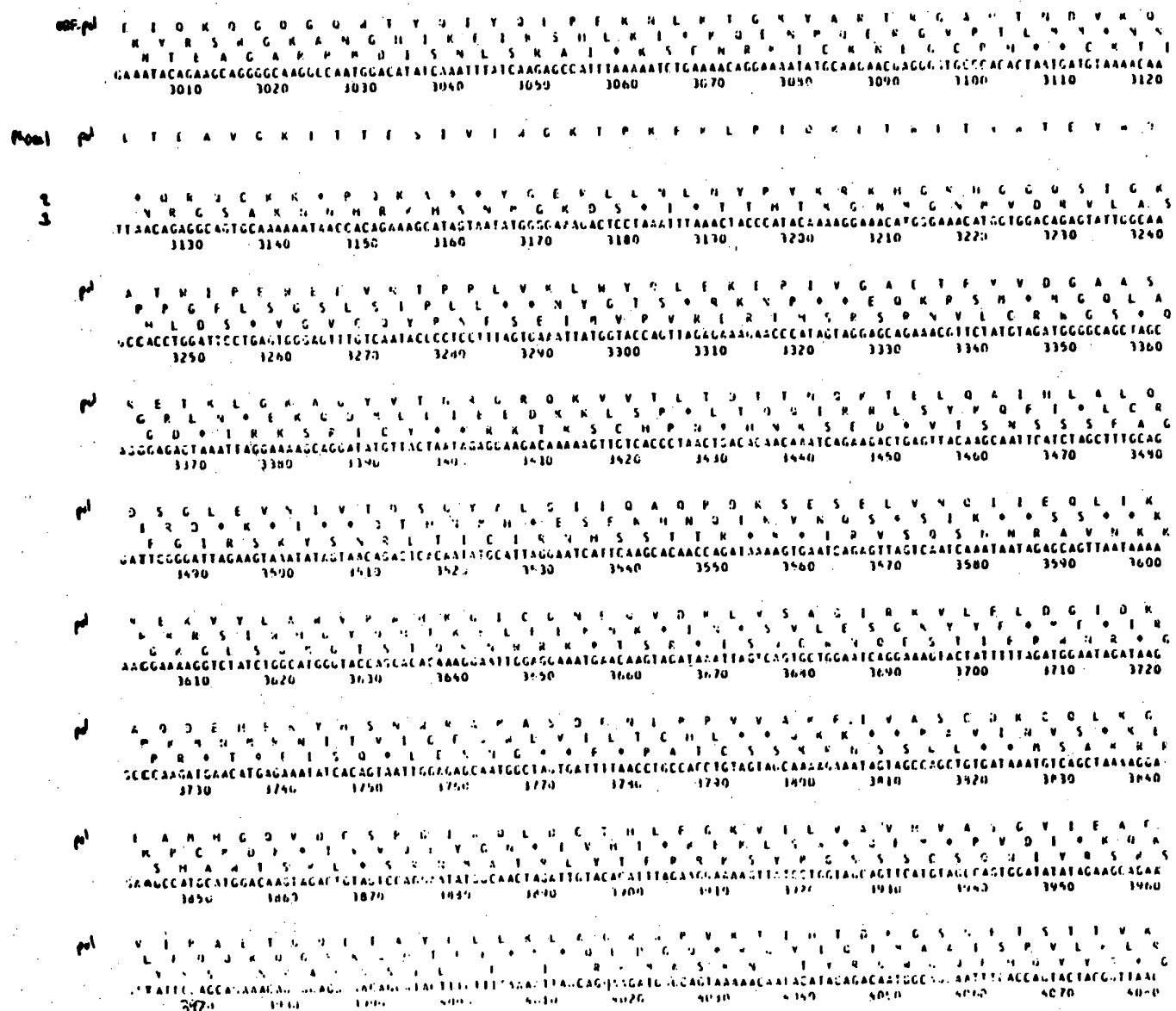


Fig 7

Fig. 8

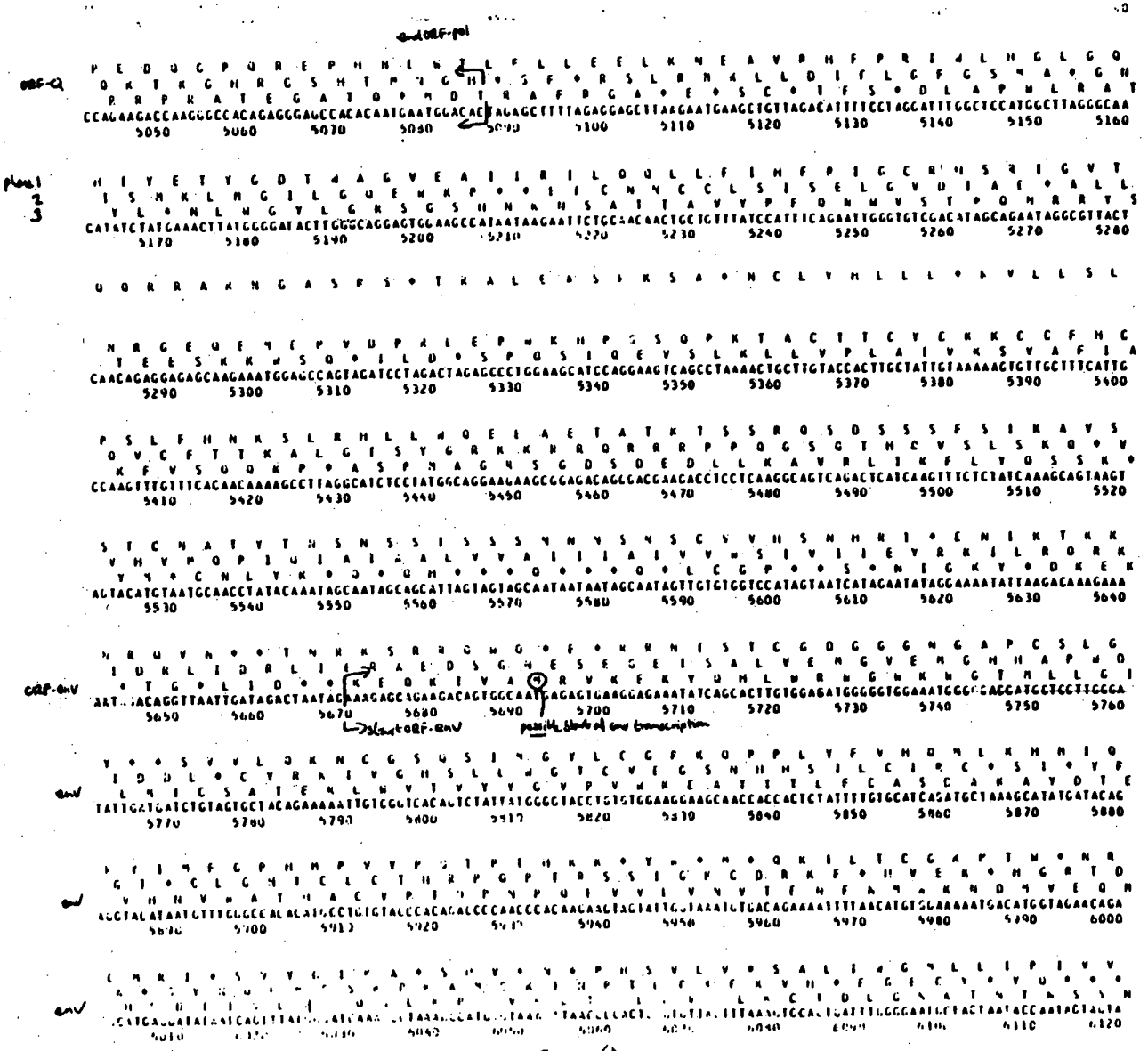


Fig 5

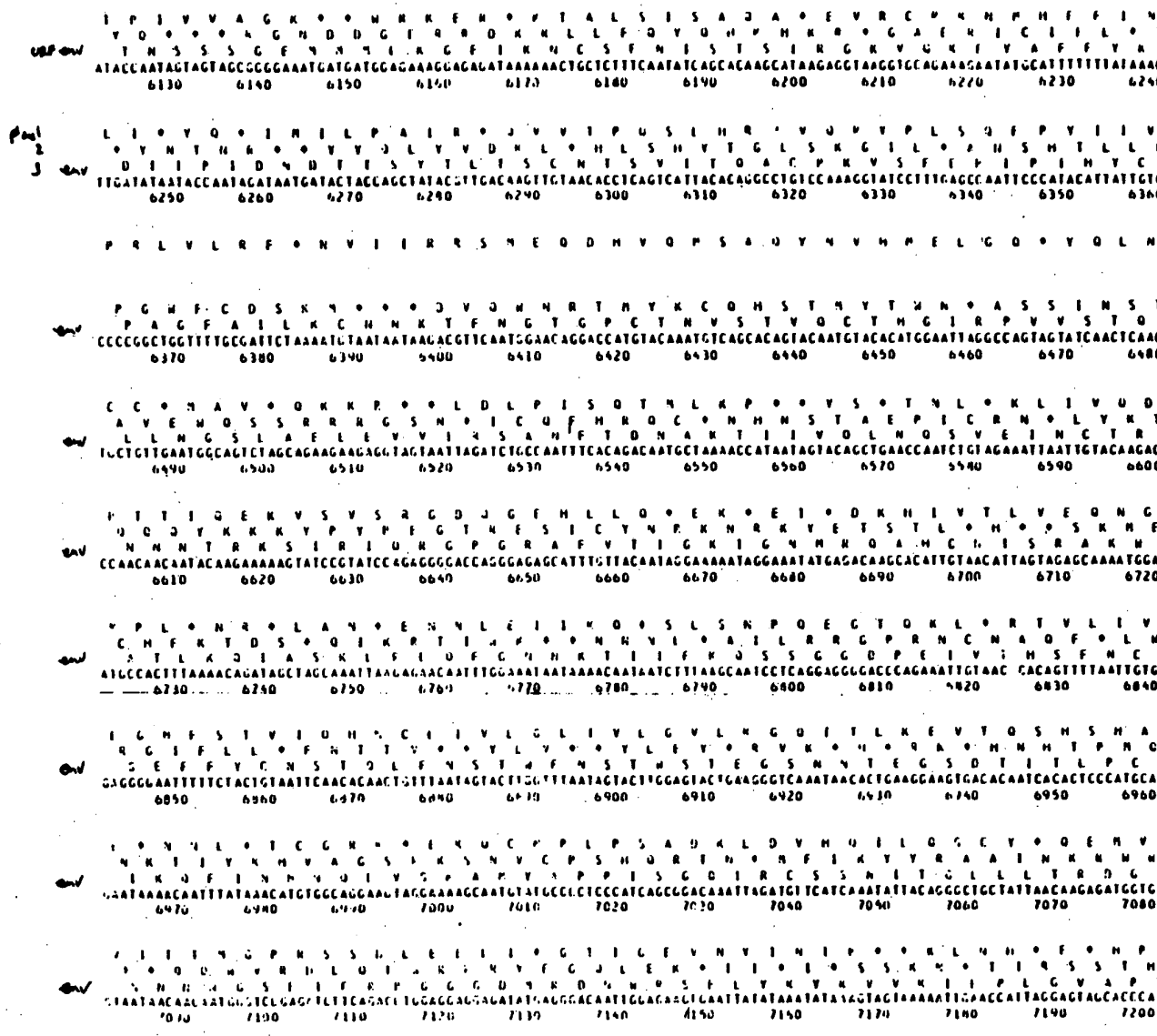


Fig 10

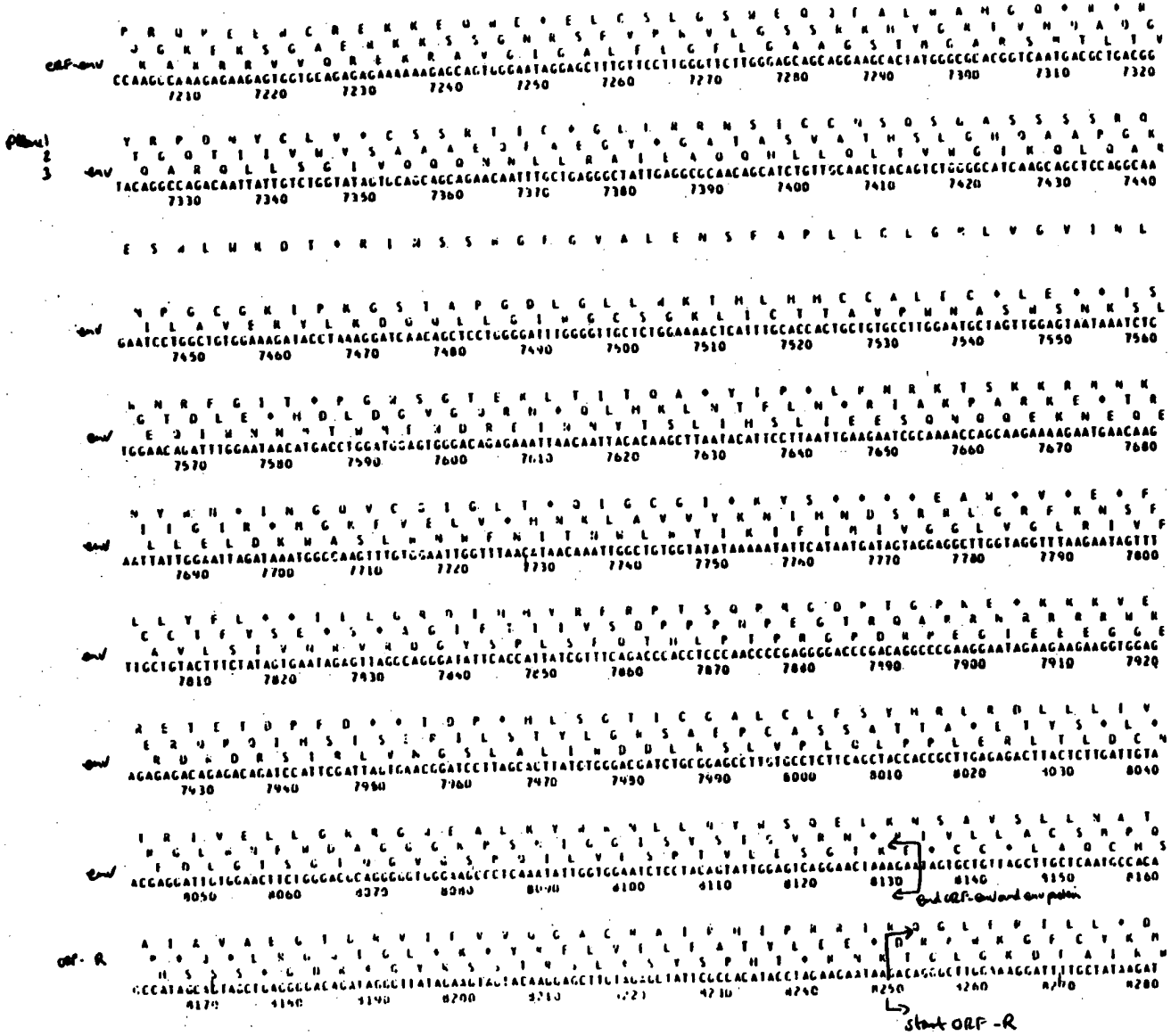


Fig 11

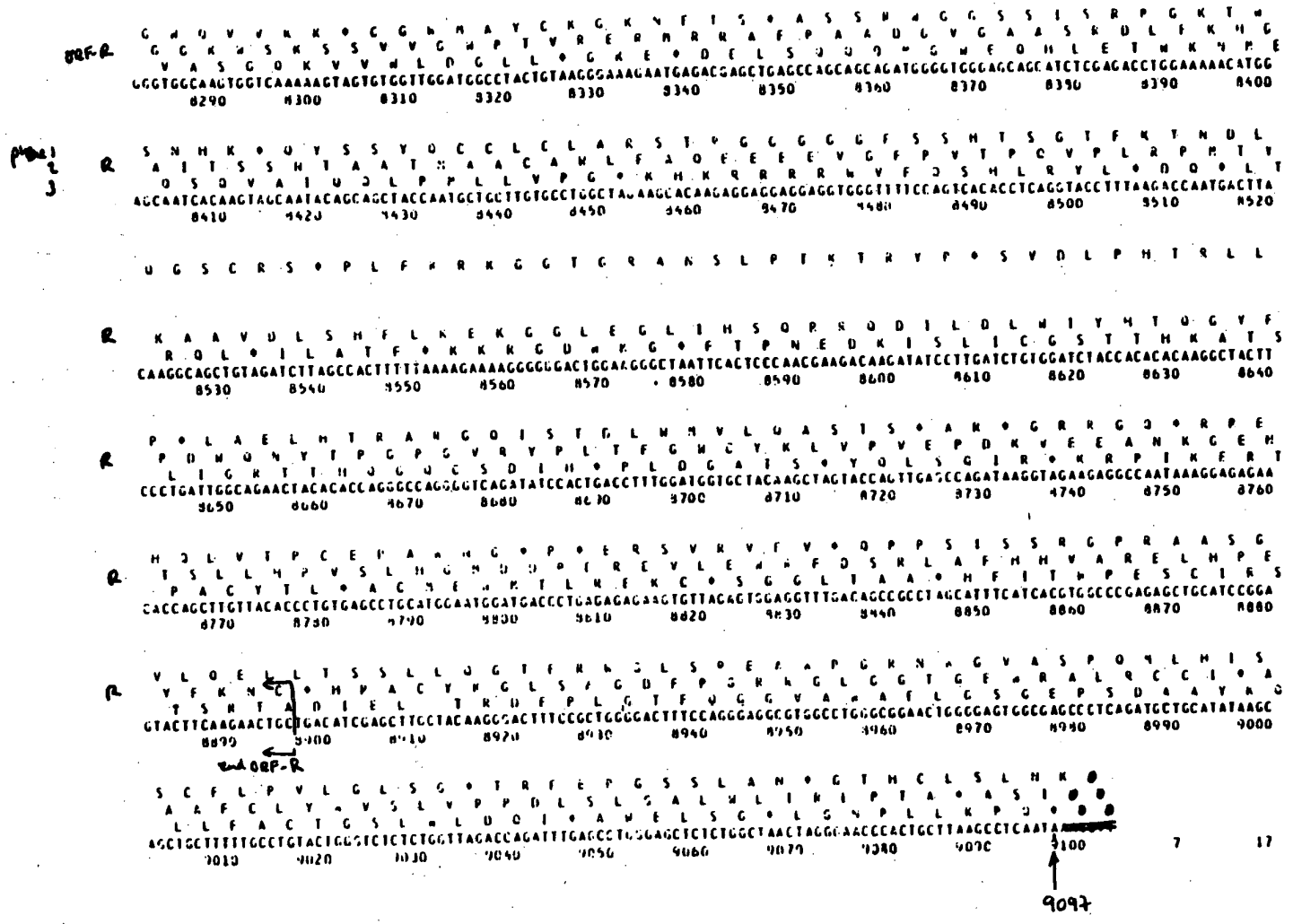


fig 12

34

08/179920

Fig 13

771248

08/177920

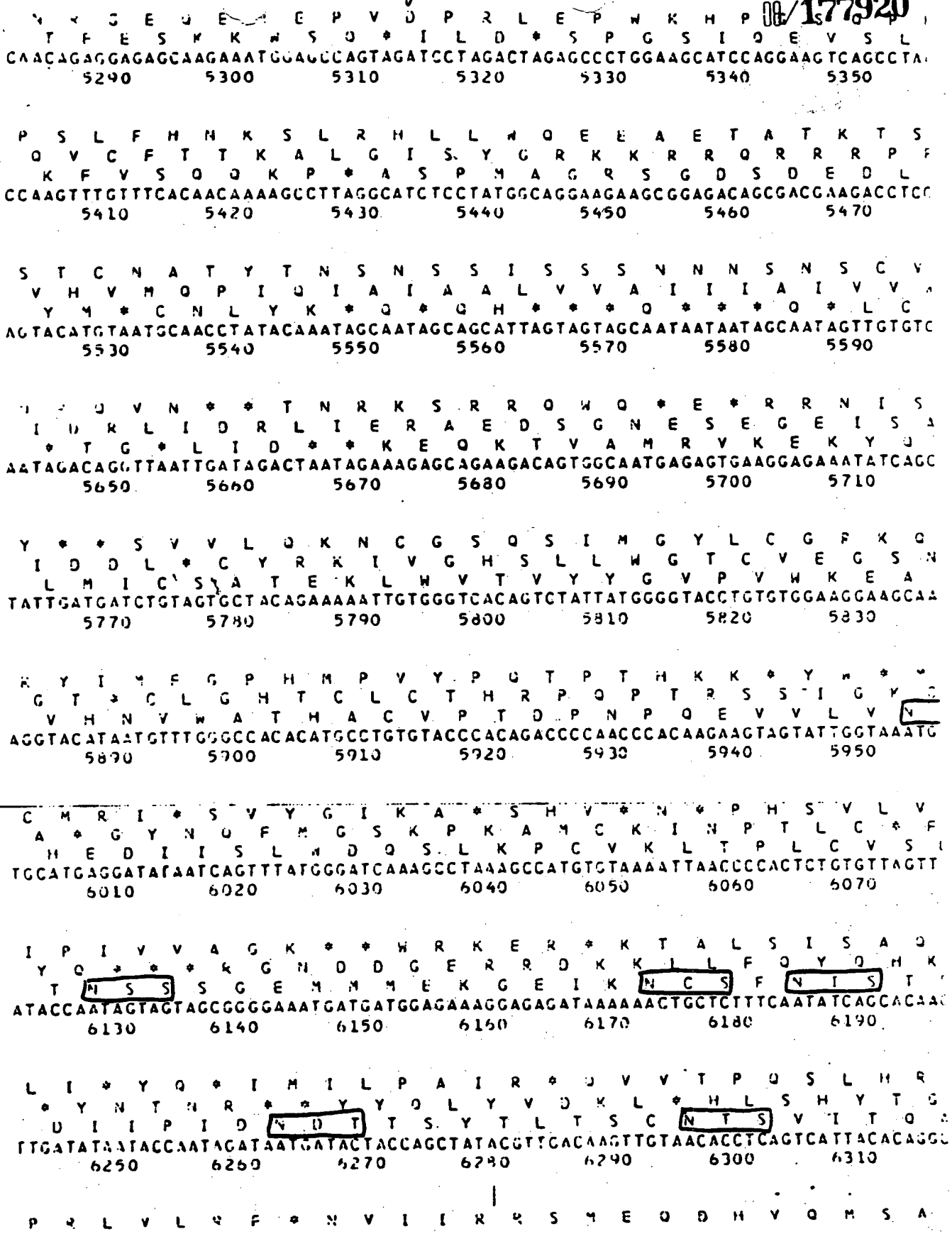


Fig 14

08/177920

```
  G  S  Q  P  K  T  C  T  T  C  Y  C  K  K  C  C  F  H  C
  Q  E  V  S  L  K  L  L  V  P  L  A  I  V  K  S  V  A  F  I  A
AGGAAGTCAGCCTAAAACTGCTTGTACCACTTGCTATTGTAAAAAGTGTTGCTTTCATTG
      5350       5360       5370       5380       5390       5400

  A  T  K  T  S  S  R  Q  S  D  S  S  S  F  S  I  K  A  V  S
  R  R  R  P  P  Q  G  S  C  T  H  C  V  S  L  S  K  Q  *  V
  S  D  E  D  L  L  K  A  V  R  L  I  K  F  L  Y  Q  S  S  K  *
AGCGACGAAGACCTCCTCAAGGCAGTCAGACTCATCAAGTTTCTCTATCAAAGCAGTAAGT
      5470       5480       5490       5500       5510       5520

  S  N  S  C  V  V  H  S  N  H  R  I  *  E  N  I  K  T  K  K
  A  I  V  V  W  S  I  V  I  I  E  Y  R  K  I  L  R  Q  R  K
  *  Q  *  L  C  G  P  *  *  S  *  N  I  G  K  Y  *  D  K  E  K
AAGCAATAGTTGTGTGGTCCATAGTAATCATAGAATATAGGAAAATATTAAGACAAAGAAA
      5590       5600       5610       5620       5630       5640

  R  R  N  I  S  T  C  G  D  G  G  G  N  G  A  P  C  S  L  G
  G  E  I  S  A  L  V  E  M  G  V  E  M  G  H  H  A  P  W  D
  K  E  K  Y  Q  H  L  W  R  W  G  W  K  W  G  T  M  L  L  G  I
AAGGAGAAATATCAGCACTTGTGGAGATGGGGGTGGAAATGGGGCACCATGCTCCTTGGGA
      5710       5720       5730       5740       5750       5760

  C  G  P  K  Q  P  P  L  Y  F  V  H  Q  M  L  K  H  M  I  Q
  V  E  G  S  N  H  H  S  I  L  C  I  R  C  *  S  I  *  Y  R
  V  W  K  E  A  T  T  T  L  F  C  A  S  D  A  K  A  Y  D  T  E
TGTGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATATGATACAG
      5830       5840       5850       5860       5870       5880

  *  Y  *  *  M  *  Q  K  I  L  T  C  G  K  M  T  W  *  N  R
  S  I  G  K  C  D  R  K  F  *  H  V  E  K  *  H  G  R  T  D
  V  V  L  V  N  V  T  E  N  F  N  M  W  K  N  D  M  V  E  Q  M
TAGTATTGGTAAATGTGACAGAAAATTTTAACATGTGGAAAAATGACATGGTAGAACAGA
      5950       5960       5970       5980       5990       6000

  H  S  V  L  V  *  S  A  L  I  W  G  W  L  L  I  P  I  V  V
  T  L  C  *  F  K  V  H  *  F  G  E  C  Y  *  Y  Q  *  *  *
  Q  L  C  V  S  L  K  C  T  D  L  G  N  A  T  N  T  N  S  S  N
CACTCTGTGTTAGTTTAAAGTGCACTGATTTGGGGAATGCTACTAATACCAATAGTAGTA
      6070       6080       6090       6100       6110       6120

  S  I  S  A  Q  A  *  E  V  R  C  P  K  N  M  H  F  F  I  N
  Q  Y  Q  H  K  H  K  R  *  G  A  E  R  I  C  I  F  L  *  T
  =  N  I  S  T  S  I  R  G  K  V  Q  K  E  Y  A  F  F  Y  K  L
TCAATATCAGCACAAGCATAAGAGGTAAGGTGCAGAAAGAATATGCATTTTTTTATAAAC
      6190       6200       6210       6220       6230       6240

  Q  S  L  H  R  P  V  Q  R  Y  P  L  S  Q  F  P  Y  I  I  V
  S  H  Y  T  G  L  S  K  G  I  L  *  A  N  S  H  T  L  L  C
  S  V  I  T  Q  A  C  P  K  V  S  F  E  P  I  P  I  H  Y  C  A
CAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACATTATTGTG
      6310       6320       6330       6340       6350       6360

  V  Q  R  S  A  Q  Y  N  V  H  M  F  L  G  Q  *  Y  Q  L  N
```

360

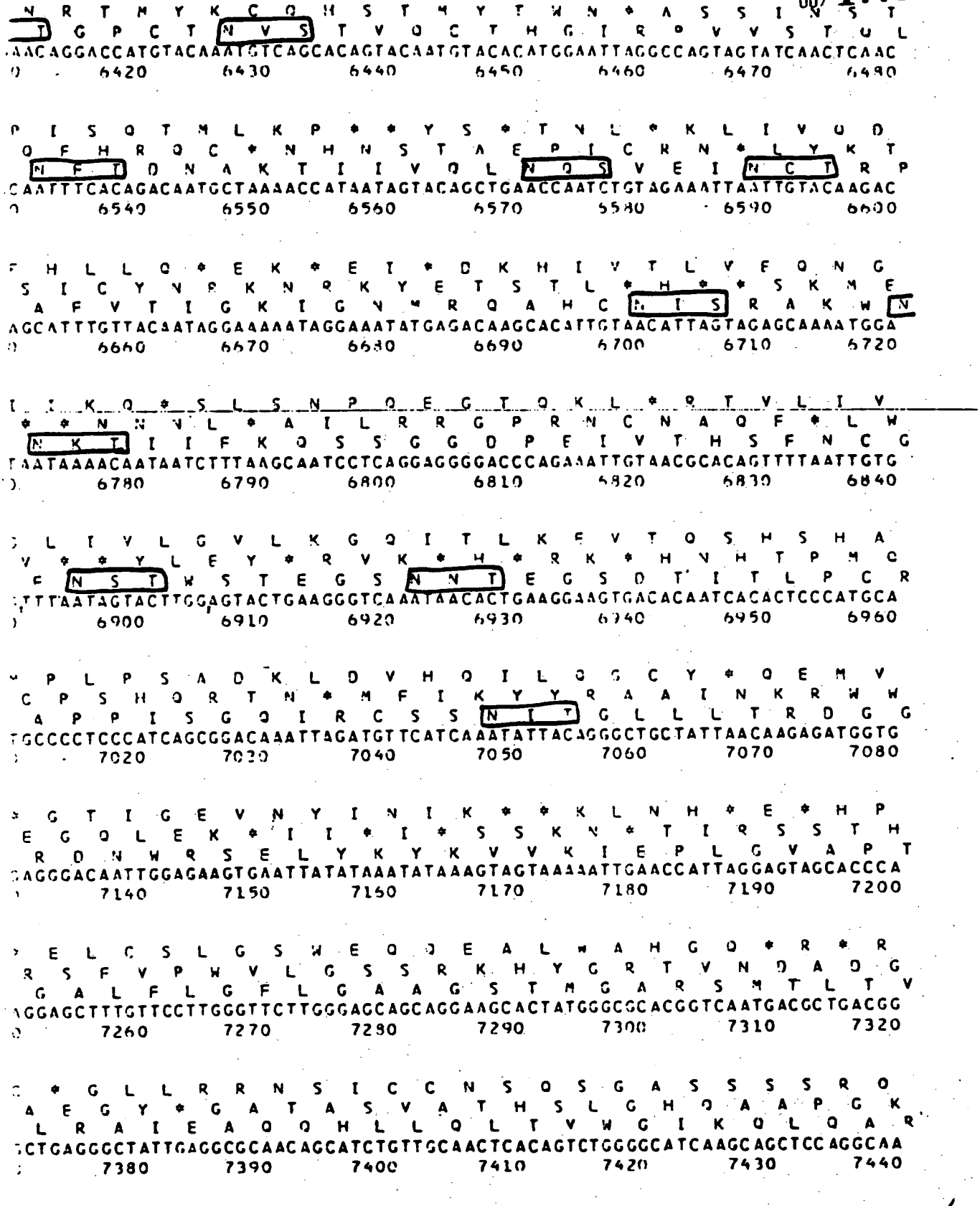
Fig 15


6370 6380 6390 6400 6410 6420 6430
6490 6500 6510 6520 6530 6540 6550
6610 6620 6630 6640 6650 6660 6670
6730 6740 6750 6760 6770 6780 6790
6850 6860 6870 6880 6890 6900 6910
6970 6980 6990 7000 7010 7020 7030
7090 7100 7110 7120 7130 7140 7150
7210 7220 7230 7240 7250 7260 7270
7330 7340 7350 7360 7370 7380 7390

771244

Fig. 16

08/177920



36

Fig 17

771248

08/177920

```
  N  P  G  C  G  K  I  P  K  G  S  T  A  P  G  D  L  G  L  K  H
  I  L  A  V  E  R  Y  L  K  D  Q  Q  L  L  G  I  W  G  C  S  G  K  L  I
GAATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCAT
      7450       7460       7470       7480       7490       7500       7510

  W  N  R  F  G  I  T  *  P  G  W  S  G  T  E  K  L  T  I  T  Q  A  *  Y
  G  T  D  L  E  *  H  D  L  D  G  V  G  Q  R  N  *  Q  L  H  K  L  N  T
  E  Q  I  W  N  N  M  T  W  M  E  W  D  R  E  I  N  N  Y  T  S  L  I  H
TGGAACAGATTTGGAATAACATGACCTGGATGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACA
      7570       7580       7590       7600       7610       7620       7630

  N  Y  W  N  *  I  N  G  Q  V  C  G  I  G  L  T  *  Q  I  G  C  G  I  *
  I  I  G  I  R  *  M  G  K  F  V  E  L  V  *  H  N  K  L  A  V  V  Y  K
  L  L  E  L  D  K  W  A  S  L  W  N  W  F  N  I  T  N  W  L  W  Y  I  K
AATTATTGGAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAA
      7690       7700       7710       7720       7730       7740       7750

  L  L  Y  F  L  *  *  I  E  L  G  R  D  I  H  H  Y  R  F  R  P  T  S  Q
  C  C  T  F  Y  S  E  *  S  *  A  G  I  F  T  I  I  V  S  D  P  P  P  N
  A  V  L  S  I  V  N  R  V  R  Q  G  Y  S  P  L  S  F  Q  T  H  L  P  T
TTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAAC
      7810       7820       7830       7840       7850       7860       7870

  R  E  T  E  T  D  P  F  D  *  *  T  D  P  *  H  L  S  G  T  I  C  G  A
  E  R  Q  R  Q  I  H  S  I  S  E  R  I  L  S  T  Y  L  G  R  S  A  E  P
  R  D  R  D  R  S  I  R  L  V  N  G  S  L  A  L  I  W  D  D  L  R  S  L
AGAGAGACAGAGACAGATCCATTCGATTAGTGAACGGATCCTTAGCACTTATCTGGGACGATCTGCGGAGCCT
      7930       7940       7950       7960       7970       7980       7990

  T  R  I  V  E  L  L  G  R  R  G  W  E  A  L  K  Y  W  W  N  L  L  Q  Y
  R  G  L  W  N  F  W  D  A  G  G  G  K  P  S  N  I  G  G  I  S  Y  S  I
  E  D  C  G  T  S  G  T  Q  G  V  G  S  P  Q  I  L  V  E  S  P  T  V  L
ACGAGGATTGTGGAACTTCTGGGACGCAGGGGGTGGGAAGCCCTCAAATATTGGTGGAATCTCCTACAGTATT
      8050       8060       8070       8080       8090       8100       8110

  A  I  A  V  A  E  G  T  D  R  V  I  E  V  V  Q  G  A  C  R  A  I  R  H  I
  P  *  Q  *  L  R  G  Q  I  G  L  *  K  *  Y  K  E  L  V  E  L  F  A  T
  H  S  S  S  *  G  D  R  *  G  Y  R  S  S  T  R  S  L  *  S  Y  S  P  H
GCCATAGCAGTAGCTGAGGGGACAGATAGGGTTATAGAAGTAGTACAAGGAGCTTGTAGAGCTATTCGCCACAT
      8170       8180       8190       8200       8210       8220       8230

  G  W  Q  V  V  K  K  *  C  G  W  M  A  Y  C  K  G  K  N  E  T  S  *  A  S
  G  G  K  W  S  K  S  S  V  V  G  W  P  T  V  R  E  R  M  R  R  A  E  P
  V  A  S  G  Q  K  V  V  W  L  D  G  L  L  *  G  K  E  *  D  E  L  S  Q
GGGTGGCAAGTGGTCAAAAAGTAGTGTGGTTGGATGGCCTACTGTAAGGGAAAGAATGAGACGAGCTGAGCCAG
      8290       8300       8310       8320       8330       8340       8350

  S  N  H  K  *  Q  Y  S  S  Y  Q  C  C  L  C  L  A  R  S  T  R  G  G  G  G
  A  I  T  S  S  N  T  A  A  T  N  A  A  C  A  W  L  F  A  Q  E  E  E  E
  Q  S  Q  V  A  I  Q  Q  L  P  M  L  L  V  P  G  *  K  H  K  R  R  R  R
AGCAATCACAAGTAGCAATACAGCAGCTACCAATGCTGCTTGTGCCTGGCTAGAAGCACAAGAGGAGGAGGAGG
      8410       8420       8430       8440       8450       8460       8470

  Q  G  S  C  R  S  *  P  L  F  K  R  K  G  G  T  G  R  A  N  S  L  P  T  K
```

39

Wind

Fig 18

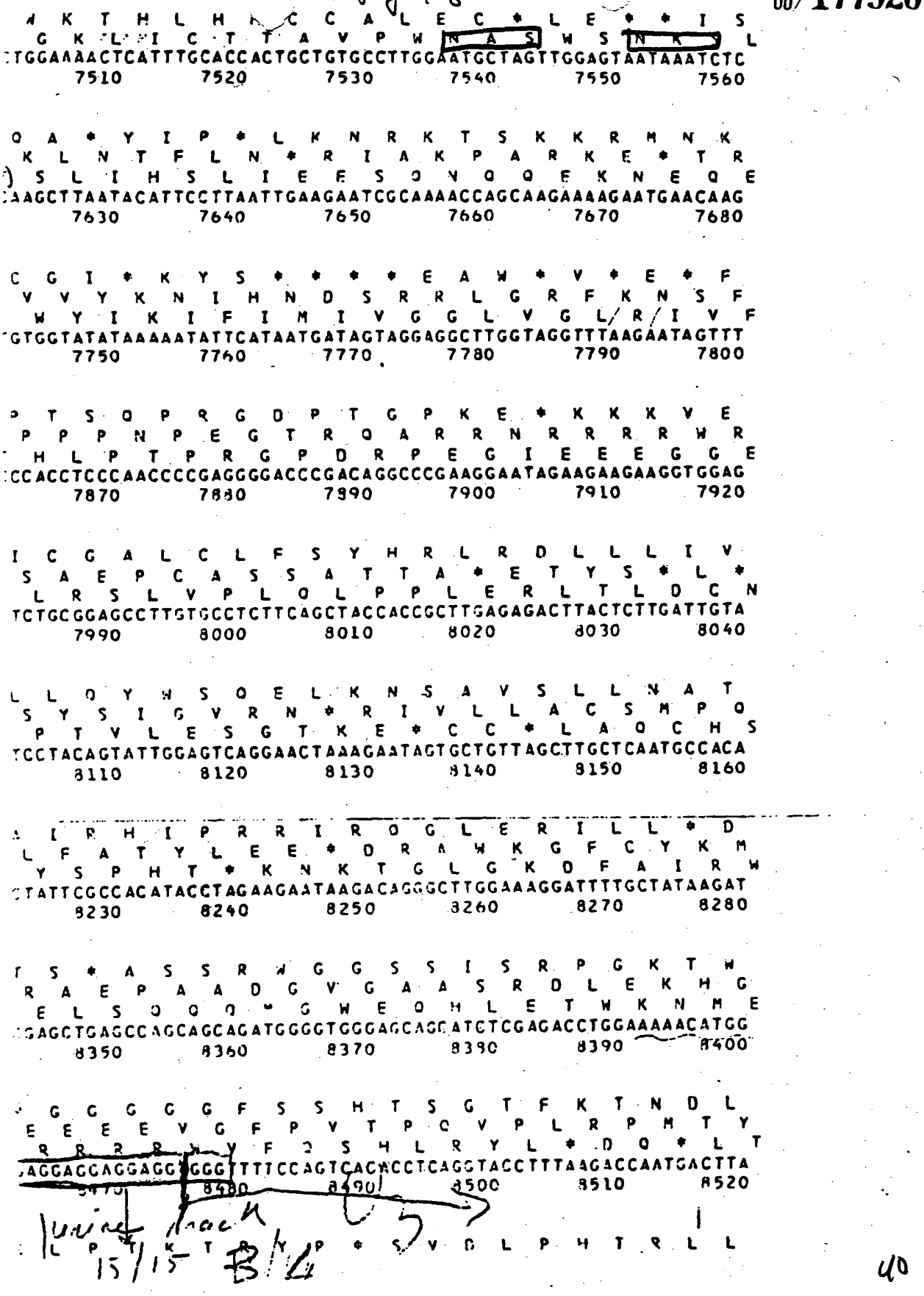

7510 7520 7530 7540 7550 7560
7630 7640 7650 7660 7670 7680
7750 7760 7770 7780 7790 7800
7870 7880 7890 7900 7910 7920
7990 8000 8010 8020 8030 8040
8110 8120 8130 8140 8150 8160
8230 8240 8250 8260 8270 8280
8350 8360 8370 8380 8390 8400
8470 8480 8490 8500 8510 8520
using Mac
15/15 B/L

Fig 19

```
        10         20         30         40         50         60
AAGCTTGCCT TGAGTGCTTC AAGTAGTGTG TGCCCGTCTG TTGTGTGACT CTGGTAACTA

        70         80         90        100        110        120
GAGATCCCTC AGACCCTTTT AGTCAGTGTG GAAAATCTCT AGCAGTGGCG CCCGAACAGG

       130        140        150        160        170        180
GACTTGAAAG CGAAAGGGAA ACCAGAGGAG CTCTCTCGAC GCAGGACTCG GCTTGCTGAA

       190        200        210        220        230        240
GCGCGCACGG CAAGAGGCGA GGGGAGGCGA CTGGTGAGTA CGCCAAAAAT TTTGACTAGC

       250        260        270        280        290        300
GGAGGCTAGA AGGAGAGAGA TGGGTGCGAG AGCGTCAGTA TTAAGCGGGG GAGAATTAGA

       310        320        330        340        350        360
TCGATGGGAA AAAATTCGGT TAAGGCCAGG GGGAAAGAAA AAATATAAAT TAAAACATAT

       370        380        390        400        410        420
AGTATGGGCA AGCAGGGAGC TAGAACGATT CGCTGTTAAT CCTGGCCTGT TAGAAACATC

       430        440        450        460        470        480
AGAAGGCTGT AGACAAATAC TGGGACAGCT ACAACCATCC CTTCAGACAG GATCAGAAGA

       490        500        510        520        530        540
ACTTAGATCA TTATATAATA CAGTAGCAAC CCTCTATTGT GTGCATCAAA GGATAGAGAT

       550        560        570        580        590        600
AAAAGACACC AAGGAAGCTT TAGACAAGAT AGAGGAAGAG CAAAACAAAA GTAAGAAAAA

       610        620        630        640        650        660
AGCACAGCAA GCAGCAGCTG ACACAGGACA CAGCAGCCAG GTCAGCCAAA ATTACCCTAT

       670        680        690        700        710        720
AGTGCAGAAC ATCCAGGGGC AAATGGTACA TCAGGCCATA TCACCTAGAA CTTTAAATGC

       730        740        750        760        770        780
ATGGGTAAAA GTAGTAGAAG AGAAGGCTTT CAGCCCAGAA GTGATACCCA TGTTTTCAGC

       790        800        810        820        830        840
ATTATCAGAA GGAGCCACCC CACAAGATTT AAACACCATG CTAAACACAG TGGGGGGACA

       850        860        870        880        890        900
TCAAGCAGCC ATGCAAATGT TAAAAGAGAC CATCAATGAG GAAGCTGCAG AATGGGATAG

       910        920        930        940        950        960
AGTGCATCCA GTGCATGCAG GGCCTATTGC ACCAGGCCAG ATGAGAGAAC CAAGGGGAAG

       970        980        990       1000       1010       1020
TGACATAGCA GGAACTACTA GTACCCTTCA GGAACAAATA GGATGGATGA CAAATAATCC

      1030       1040       1050       1060       1070       1080
ACCTATCCCA GTAGGAGAAA TTTATAAAAG ATGGATAATC CTGGGATTAA ATAAAATAGT

      1090       1100       1110       1120       1130       1140
```

Fig 20

```
AATAATGTAT AGCCCTACCA GCATTCTGGA CATAAGACAA GGACCAAAAG AACCCTTTAG

      1150        1160        1170        1180        1190        1200
AGACTATGTA GACCGGTTCT ATAAAACTCT AAGAGCCGAG CAAGCTTCAC AGGAGGTAAA

      1210        1220        1230        1240        1250        1260
AAATTGGATG ACAGAAACCT TGTTGGTCCA AAATGCGAAC CCAGATTGTA AGACTATTTT

      1270        1280        1290        1300        1310        1320
AAAAGCATTG GGACCAGCAG CTACACTAGA AGAAATGATG ACAGCATGTC AGGGAGTGGG

      1330        1340        1350        1360        1370        1380
AGGACCCGGC CATAAGGCAA GAGTTTTGGC TGAAGCAATG AGCCAAGTAA CAAATTCAGC

      1390        1400        1410        1420        1430        1440
TACCATAATG ATGCAAAGAG GCAATTTTAG GAACCAAAGA AAGATTGTTA AGTGTTTCAA

      1450        1460        1470        1480        1490        1500
TTGTGGCAAA GAAGGGCACA TAGCCAGAAA TTGCAGGGCC CCTAGGAAAA AGGGCTGTTG

      1510        1520        1530        1540        1550        1560
GAAATGTGGA AAGGAAGGAC ACCAAATGAA AGATTGTACT GAGAGACAGG CTAATTTTTT

      1570        1580        1590        1600        1610        1620
AGGGAAGATC TGGCCTTCCT ACAAGGGAAG GCCAGGGAAT TTTCTTCAGA GCAGACCAGA

      1630        1640        1650        1660        1670        1680
GCCAACAGCC CCACCAGAAG AGAGCTTCAG GTCTGGGGTA GAGACAACAA CTCCCTCTCA

      1690        1700        1710        1720        1730        1740
GAAGCAGGAG CCGATAGACA AGGAACTGTA TCCTTTAACT TCCCTCAGAT CACTCTTTGG

      1750        1760        1770        1780        1790        1800
CAACGACCCC TCGTCACAAT AAAGATAGGG GGGCAACTAA AGGAAGCTCT ATTAGATACA

      1810        1820        1830        1840        1850        1860
GGAGCAGATG ATACAGTATT AGAAGAAATG AGTTTGCCAG GAAGATGGAA ACCAAAAATG

      1870        1880        1890        1900        1910        1920
ATAGGGGGAA TTGGAGGTTT TATCAAAGTA AGACAGTATG ATCAGATACT CATAGAAATC

      1930        1940        1950        1960        1970        1980
TGTGGACATA AAGCTATAGG TACAGTATTA GTAGGACCTA CACCTGTCAA CATAATTGGA

      1990        2000        2010        2020        2030        2040
AGAAATCTGT TGACTCAGAT TGGTTGCACT TTAAATTTTC CCATTAGTCC TATTGAAACT

      2050        2060        2070        2080        2090        2100
GTACCAGTAA AATTAAAGCC AGGAATGGAT GGCCCAAAAG TTAAACAATG GCCATTGACA

      2110        2120        2130        2140        2150        2160
GAAGAAAAAA TAAAAGCATT AGTAGAAATT TGTACAGAAA TGGAAAAGGA AGGGAAAATT

      2170        2180        2190        2200        2210        2220
TCAAAAATTG GGCCTGAAAA TCCATACAAT ACTCCAGTAT TTGCCATAAA GAAAAAAGAC

      2230        2240        2250        2260        2270        2280
AGTACTAAAT GGAGAAAATT AGTAGATTTC AGAGAACTTA ATAAGAGAAC TCAAGACTTC

      2290        2300        2310        2320        2330        2340
TGGGAAGTTC AATTAGGAAT ACCACATCCC GCAGGGTTAA AAAAGAAAAA ATCAGTAACA

      2350        2360        2370        2380        2390        2400
```

42

Fig 21

GTGCTGGATG TGGGTGATGC ATATTTTTCA GTTCCCTTAG ATGAAGACTT CAGGAAGTAT

2410 2420 2430 2440 2450 2460
ACTGCATTTA CCATACCTAG TATAAACAAT GAGACACCAG GGATTAGATA TCAGTACAAT

2470 2480 2490 2500 2510 2520
GTGCTTCCAC AGGGATGGAA AGGATCACCA GCAATATTCC AAAGTAGCAT GACAAAAATC

2530 2540 2550 2560 2570 2580
TTAGAGCCTT TTAGAAAACA AAATCCAGAC ATAGTTATCT ATCAATACAT GGATGATTTG

2590 2600 2610 2620 2630 2640
TATGTAGGAT CTGACTTAGA AATAGGGCAG CATAGAACAA AAATAGAGGA GCTGAGACAA

2650 2660 2670 2680 2690 2700
CATCTGTTGA GGTGGGGACT TACCACACCA GACAAAAAAC ATCAGAAAGA ACCTCCATTC

2710 2720 2730 2740 2750 2760
CTTTGGATGG GTTATGAACT CCATCCTGAT AAATGGACAG TACAGCCTAT AGTGCTGCCA

2770 2780 2790 2800 2810 2820
GAAAAAGACA GCTGGACTGT CAATGACATA CAGAAGTTAG TGGGAAAATT GAATTGGGCA

2830 2840 2850 2860 2870 2880
AGTCAGATTT ACCCAGGGAT TAAAGTAAGG CAATTATGTA AACTCCTTAG AGGAACCAAA

2890 2900 2910 2920 2930 2940
GCACTAACAG AAGTAATACC ACTAACAGAA GAAGCAGAGC TAGAACTGGC AGAAAACAGA

2950 2960 2970 2980 2990 3000
GAGATTCTAA AAGAACCAGT ACATGGAGTG TATTATGACC CATCAAAAGA CTTAATAGCA

3010 3020 3030 3040 3050 3060
GAAATACAGA AGCAGGGGCA AGGCCAATGG ACATATCAAA TTTATCAAGA GCCATTTAAA

3070 3080 3090 3100 3110 3120
AATCTGAAAA CAGGAAAATA TGCAAGAACG AGGGGTGCCC ACACTAATGA TGTAAAACAA

3130 3140 3150 3160 3170 3180
TTAACAGAGG CAGTGCAAAA AATAACCACA GAAAGCATAG TAATATGGGG AAAGACTCCT

3190 3200 3210 3220 3230 3240
AAATTTAAAC TACCCATACA AAAGGAAACA TGGGAAACAT GGTGGACAGA GTATTGGCAA

3250 3260 3270 3280 3290 3300
GCCACCTGGA TTCCTGAGTG GGAGTTTGTC AATACCCCTC CTTTAGTGAA ATTATGGTAC

3310 3320 3330 3340 3350 3360
CAGTTAGAGA AAGAACCCAT AGTAGGAGCA GAAACGTTCT ATGTAGATGG GGCAGCTAGC

3370 3380 3390 3400 3410 3420
AGGGAGACTA AATTAGGAAA AGCAGGATAT GTTACTAATA GAGGAAGACA AAAAGTTGTC

3430 3440 3450 3460 3470 3480
ACCCTAACTG ACACAACAAA TCAGAAGACT GAGTTACAAG CAATTCATCT AGCTTTGCAG

3490 3500 3510 3520 3530 3540
GATTCGGGAT TAGAAGTAAA TATAGTAACA GACTCACAAT ATGCATTAGG AATCATTCAA

3550 3560 3570 3580 3590 3600
GCACAACCAG ATAAAAGTGA ATCAGAGTTA GTCAATCAAA TAATAGAGCA GTTAATAAAA

43

3610 3620 3630 3640 3650 3660

Fig 22

```
AAGAAAAA   TCTATCTGGC ATGGGTACCA GCACACAAAG GAATTGGAGG AAATGAACAA

      3670        3680        3690        3700        3710        3720
GTAGATAAAT TAGTCAGTGC TGCAATCAGG AAAGTACTAT TTTTAGATGG AATAGATAAG

      3730        3740        3750        3760        3770        3780
GCCCAAGATG AACATGAGAA ATATCACAGT AATTGGAGAG CAATGGCTAG TGATTTTAAC

      3790        3800        3810        3820        3830        3840
CTGCCACCTG TAGTAGCAAA AGAAATAGTA GCCAGCTGTG ATAAATGTCA GCTAAAAGGA

      3850        3860        3870        3880        3890        3900
GAAGCCATGC ATGGACAAGT AGACTGTAGT CCAGGAATAT GGCAACTAGA TTGTACACAT

      3910        3920        3930        3940        3950        3960
TTAGAAGGAA AAGTTATCCT GGTAGCAGTT CATGTAGCCA GTGGATATAT AGAAGCAGAA

      3970        3980        3990        4000        4010        4020
GTTATTCCAG CAGAAACAGG GCAGGAAACA GCATACTTTC TTTTAAAATT AGCAGGAAGA

      4030        4040        4050        4060        4070        4080
TGGCCAGTAA AAACAATACA TACAGACAAT GGCAGCAATT TCACCAGTAC TACGGTTAAG

      4090        4100        4110        4120        4130        4140
GCCGCCTGTT GGTGGGCGGG AATCAAGCAG GAATTTGGAA TTCCCTACAA TCCCCAAAGT

      4150        4160        4170        4180        4190        4200
CAAGGAGTAG TAGAATCTAT GAATAAAGAA TTAAAGAAAA TTATAGGCCA GGTAAGAGAT

      4210        4220        4230        4240        4250        4260
CAGGCTGAAC ATCTTAAGAC AGCAGTACAA ATGGCAGTAT TCATCCACAA TTTTAAAAGA

      4270        4280        4290        4300        4310        4320
AAAGGGGGGA TTGGGGGGTA CAGTGCAGGG GAAAGAATAG TAGACATAAT AGCAACAGAC

      4330        4340        4350        4360        4370        4380
ATACAAACTA AAGAATTACA AAAACAAATT ACAAAAATTC AAAATTTTCG GGTTTATTAC

      4390        4400        4410        4420        4430        4440
AGGGACAGCA GAGATCCACT TTGGAAAGGA CCAGCAAAGC TCCTCTGGAA AGGTGAAGGG

      4450        4460        4470        4480        4490        4500
GCAGTAGTAA TACAAGATAA TAGTGACATA AAAGTAGTGC CAAGAAGAAA AGCAAAGATC

      4510        4520        4530        4540        4550        4560
ATTAGGGATT ATGGAAAACA GATGGCAGGT GATGATTGTG TGGCAAGTAG ACAGGATGAG

      4570        4580        4590        4600        4610        4620
GATTAGAACA TGGAAAAGTT TAGTAAAACA CCATATGTAT GTTTCAGGGA AAGCTAGGGG

      4630        4640        4650        4660        4670        4680
ATGGTTTTAT AGACATCACT ATGAAAGCCC TCATCCAAGA ATAAGTTCAG AAGTACACAT

      4690        4700        4710        4720        4730        4740
CCCACTAGGG GATGCTAGAT TGGTAATAAC AACATATTGG GGTCTGCATA CAGGAGAAAG

      4750        4760        4770        4780        4790        4800
AGACTGGCAT CTGGGTCAGG GAGTCTCCAT AGAATGGAGG AAAAAGAGAT ATAGCACACA

      4810        4820        4830        4840        4850        4860
AGTAGACCCT GAACTAGCAG ACCAACTAAT TCATCTGTAT TACTTTGACT GTTTTTCAGA

      4870        4880        4890        4900        4910        4920
```

44

CTCTGCTATA AGAAAGGCCT TATTAGGACA TATAGTTAGC CCTAGGTGTG AATATCAAGC

4930 4940 4950 4960 4970 4980
AGGACATAAC AAGGTAGGAT CTCTACAATA CTTGGCACTA GCAGCATTAA TAACACCAAA

4990 5000 5010 5020 5030 5040
AAAGATAAAG CCACCTTTGC CTAGTGTTAC GAAACTGACA GAGGATAGAT GGAACAAGCC

5050 5060 5070 5080 5090 5100
CCAGAAGACC AAGGGCCACA GAGGGAGCCA CACAATGAAT GGACACTAGA GCTTTTAGAG

5110 5120 5130 5140 5150 5160
GAGCTTAAGA ATGAAGCTGT TAGACATTTT CCTAGGATTT GGCTCCATGG CTTAGGGCAA

5170 5180 5190 5200 5210 5220
CATATCTATG AAACTTATGG GGATACTTGG GCAGGAGTGG AAGCCATAAT AAGAATTCTG

5230 5240 5250 5260 5270 5280
CAACAACTGC TGTTTATCCA TTTCAGAATT GGGTGTCGAC ATAGCAGAAT AGGCGTTACT

5290 5300 5310 5320 5330 5340
CAACAGAGGA GAGCAAGAAA TGGAGCCAGT AGATCCTAGA CTAGAGCCCT GGAAGCATCC

5350 5360 5370 5380 5390 5400
AGGAAGTCAG CCTAAAACTG CTTGTACCAC TTGCTATTGT AAAAAGTGTT GCTTTCATTG

5410 5420 5430 5440 5450 5460
CCAAGTTTGT TTCACAACAA AAGCCTTAGG CATCTCCTAT GGCAGGAAGA AGCGGAGACA

5470 5480 5490 5500 5510 5520
GCGACGAAGA CCTCCTCAAG GCAGTCAGAC TCATCAAGTT TCTCTATCAA AGCAGTAAGT

5530 5540 5550 5560 5570 5580
AGTACATGTA ATGCAACCTA TACAAATAGC AATAGCAGCA TTAGTAGTAG CAATAATAAT

5590 5600 5610 5620 5630 5640
AGCAATAGTT GTGTGGTCCA TAGTAATCAT AGAATATAGG AAAATATTAA GACAAAGAAA

5650 5660 5670 5680 5690 5700
AATAGACAGG TTAATTGATA GACTAATAGA AAGAGCAGAA GACAGTGGCA ATGAGAGTGA

5710 5720 5730 5740 5750 5760
AGGAGAAATA TCAGCACTTG TGGAGATGGG GGTGGAAATG GGGCACCATG CTCCTTGGGA

5770 5780 5790 5800 5810 5820
TATTGATGAT CTGTAGTGCT ACAGAAAAAT TGTGGGTCAC AGTCTATTAT GGGGTACCTG

5830 5840 5850 5860 5870 5880
TGTGGAAGGA AGCAACCACC ACTCTATTTT GTGCATCAGA TGCTAAAGCA TATGATACAG

5890 5900 5910 5920 5930 5940
AGGTACATAA TGTTTGGGCC ACACATGCCT GTGTACCCAC AGACCCCAAC CCACAAGAAG

5950 5960 5970 5980 5990 6000
TAGTATTGGT AAATGTGACA GAAAATTTTA ACATGTGGAA AAATGACATG GTAGAACAGA

6010 6020 6030 6040 6050 6060
TGCATGAGGA TATAATCAGT TTATGGGATC AAAGCCTAAA GCCATGTGTA AAATTAACCC

6070 6080 6090 6100 6110 6120
CACTCTGTGT TAGTTTAAAG TGCACTGATT TGGGGAATGC TACTAATACC AATAGTAGTA

6130 6140 6150 6160 6170 6180

Fig 23

US

ATACCAATAG TAGAGCGGG GAAATGATGA TGGAGAAAGG AGAGATAAAA AACTGCTCTT

6190 6200 6210 6220 6230 6240
TCAATATCAG CACAAGCATA AGAGGTAAGG TGCAGAAAGA ATATGCATTT TTTTATAAAC

6250 6260 6270 6280 6290 6300
TTGATATAAT ACCAATAGAT AATGATACTA CCAGCTATAC GTTGACAAGT TGTAACACCT

6310 6320 6330 6340 6350 6360
CAGTCATTAC ACAGGCCTGT CCAAAGGTAT CCTTTGAGCC AATTCCCATA CATTATTGTG

6370 6380 6390 6400 6410 6420
CCCCGGCTGG TTTTGCGATT CTAAAATGTA ATAATAAGAC GTTCAATGGA ACAGGACCAT

6430 6440 6450 6460 6470 6480
GTACAAATGT CAGCACAGTA CAATGTACAC ATGGAATTAG GCCAGTAGTA TCAACTCAAC

6490 6500 6510 6520 6530 6540
TGCTGTTGAA TGGCAGTCTA GCAGAAGAAG AGGTAGTAAT TAGATCTGCC AATTTCACAG

6550 6560 6570 6580 6590 6600
ACAATGCTAA AACCATAATA GTACAGCTGA ACCAATCTGT AGAAATTAAT TGTACAAGAC

6610 6620 6630 6640 6650 6660
CCAACAACAA TACAAGAAAA AGTATCCGTA TCCAGAGGGG ACCAGGGAGA GCATTTGTTA

6670 6680 6690 6700 6710 6720
CAATAGGAAA AATAGGAAAT ATGAGACAAG CACATTGTAA CATTAGTAGA GCAAAATGGA

6730 6740 6750 6760 6770 6780
ATGCCACTTT AAAACAGATA GCTAGCAAAT TAAGAGAACA ATTTGGAAAT AATAAAACAA

6790 6800 6810 6820 6830 6840
TAATCTTTAA GCAATCCTCA GGAGGGGACC CAGAAATTGT AACGCACAGT TTTAATTGTG

6850 6860 6870 6880 6890 6900
GAGGGGAATT TTTCTACTGT AATTCAACAC AACTGTTTAA TAGTACTTGG TTTAATAGTA

6910 6920 6930 6940 6950 6960
CTTGGAGTAC TGAAGGGTCA AATAACACTG AAGGAAGTGA CACAATCACA CTCCCATGCA

6970 6980 6990 7000 7010 7020
GAATAAAACA ATTTATAAAC ATGTGGCAGG AAGTAGGAAA AGCAATGTAT GCCCCTCCCA

7030 7040 7050 7060 7070 7080
TCAGCGGACA AATTAGATGT TCATCAAATA TTACAGGGCT GCTATTAACA AGAGATGGTG

7090 7100 7110 7120 7130 7140
GTAATAACAA CAATGGGTCC GAGATCTTCA GACCTGGAGG AGGAGATATG AGGGACAATT

7150 7160 7170 7180 7190 7200
GGAGAAGTGA ATTATATAAA TATAAAGTAG TAAAAATTGA ACCATTAGGA GTAGCACCCA

7210 7220 7230 7240 7250 7260
CCAAGGCAAA GAGAAGAGTG GTGCAGAGAG AAAAAAGAGC AGTGGGAATA GGAGCTTTGT

7270 7280 7290 7300 7310 7320
TCCTTGGGTT CTTGGGAGCA GCAGGAAGCA CTATGGGCGC ACGGTCAATG ACGCTGACGG

7330 7340 7350 7360 7370 7380
TACAGGCCAG ACAATTATTG TCTGGTATAG TGCAGCAGCA GAACAATTTG CTGAGGGCTA

7390 7400 7410 7420 7430 7440

TTGAGGCGCA ACAGCATCTG TTGCAACTCA CAGTCTGGGG CATCAAGCAG CTCCAGGCAA

7450 7460 7470 7480 7490 7500
GAATCCTGGC TGTGGAAAGA TACCTAAAGG ATCAACAGCT CCTGGGGATT TGGGGTTGCT

7510 7520 7530 7540 7550 7560
CTGGAAAACT CATTTGCACC ACTGCTGTGC CTTGGAATGC TAGTTGGAGT AATAAATCTC

7570 7580 7590 7600 7610 7620
TGGAACAGAT TTGGAATAAC ATGACCTGGA TGGAGTGGGA CAGAGAAATT AACAATTACA

7630 7640 7650 7660 7670 7680
CAAGCTTAAT ACATTCCTTA ATTGAAGAAT CGCAAAACCA GCAAGAAAAG AATGAACAAG

7690 7700 7710 7720 7730 7740
AATTATTGGA ATTAGATAAA TGGGCAAGTT TGTGGAATTG GTTTAACATA ACAAATTGGC

7750 7760 7770 7780 7790 7800
TGTGGTATAT AAAAATATTC ATAATGATAG TAGGAGGCTT GGTAGGTTTA AGAATAGTTT

7810 7820 7830 7840 7850 7860
TTGCTGTACT TTCTATAGTG AATAGAGTTA GGCAGGGATA TTCACCATTA TCGTTTCAGA

7870 7880 7890 7900 7910 7920
CCCACCTCCC AACCCCGAGG GGACCCGACA GGCCCGAAGG AATAGAAGAA GAAGGTGGAG

7930 7940 7950 7960 7970 7980
AGAGAGACAG AGACAGATCC ATTCGATTAG TGAACGGATC CTTAGCACTT ATCTGGGACG

7990 8000 8010 8020 8030 8040
ATCTGCGGAG CCTTGTGCCT CTTCAGCTAC CACCGCTTGA GAGACTTACT CTTGATTGTA

8050 8060 8070 8080 8090 8100
ACGAGGATTG TGGAACTTCT GGGACGCAGG GGGTGGGAAG CCCTCAAATA TTGGTGGAAT

8110 8120 8130 8140 8150 8160
CTCCTACAGT ATTGGAGTCA GGAACTAAAG AATAGTGCTG TTAGCTTGCT CAATGCCACA

8170 8180 8190 8200 8210 8220
GCCATAGCAG TAGCTGAGGG GACAGATAGG GTTATAGAAG TAGTACAAGG AGCTTGTAGA

8230 8240 8250 8260 8270 8280
GCTATTCGCC ACATACCTAG AAGAATAAGA CAGGGCTTGG AAAGGATTTT GCTATAAGAT

8290 8300 8310 8320 8330 8340
GGGTGGCAAG TGGTCAAAAA GTAGTGTGGT TGGATGGCCT ACTGTAAGGG AAAGAATGAG

8350 8360 8370 8380 8390 8400
ACGAGCTGAG CCAGCAGCAG ATGGGGTGGG AGCAGCATCT CGAGACCTGG AAAAACATGG

8410 8420 8430 8440 8450 8460
AGCAATCACA AGTAGCAATA CAGCAGCTAC CAATGCTGCT TGTGCCTGGC TAGAAGCACA

8470 8480 8490 8500 8510 8520
AGAGGAGGAG GAGGTGGGTT TTCCAGTCAC ACCTCAGGTA CCTTTAAGAC CAATGACTTA

8530 8540 8550 8560 8570 8580
CAAGGCAGCT GTAGATCTTA GCCACTTTTT AAAAGAAAAG GGGGGACTGG AAGGGCTAAT

8590 8600 8610 8620 8630 8640
TCACTCCCAA CGAAGACAAG ATATCCTTGA TCTGTGGATC TACCACACAC AAGGCTACTT

8650 8660 8670 8680 8690 8700

```
CCCTGATTGG CAGAACTACA CACCAGGGCC AGGGGTCAGA TATCCACTGA CCTTTGGATG

      8710        8720        8730        8740        8750        8760
GTGCTACAAG CTAGTACCAG TTGAGCCAGA TAAGGTAGAA GAGGCCAATA AAGGAGAGAA

      8770        8780        8790        8800        8810        8820
CACCAGCTTG TTACACCCTG TGAGCCTGCA TGGAATGGAT GACCCTGAGA GAGAAGTGTT

      8830        8840        8850        8860        8870        8880
AGAGTGGAGG TTTGACAGCC GCCTAGCATT TCATCACGTG GCCCGAGAGC TGCATCCGGA

      8890        8900        8910        8920        8930        8940
GTACTTCAAG AACTGCTGAC ATCGAGCTTG CTACAAGGGA CTTTCCGCTG GGGACTTTCC

      8950        8960        8970        8980        8990        9000
AGGGAGGCGT GGCCTGGGCG GAACTGGGGA GTGGCGAGCC CTCAGATGCT GCATATAAGC

      9010        9020        9030        9040        9050        9060
AGCTGCTTTT TGCCTGTACT GGGTCTCTCT GGTTAGACCA GATTTGAGCC TGGGAGCTCT

      9070        9080        9090        9100           0           0
CTGGCTAACT AGGGAACCCA CTGCTTAAGC CTCAATAAAG CTT
```

Fig 9b